

# Site-specific online compressive beam codebook learning in mmWave vehicular communication

Yuyang Wang, Nitin Jonathan Myers, Nuria González-Prelcic,  
and Robert W. Heath Jr.

**Abstract**—Millimeter wave (mmWave) communication is one viable solution to support Gbps sensor data sharing in vehicular networks. The use of large antenna arrays at mmWave and high mobility in vehicular communication make it challenging to design fast beam alignment solutions. In this paper, we propose a novel framework that learns the channel angle-of-departure (AoD) statistics at a base station (BS) and uses this information to efficiently acquire channel measurements. Our framework integrates online learning for compressive sensing (CS) codebook learning and the optimized codebook is used for CS-based beam alignment. We formulate a CS matrix optimization problem based on the AoD statistics available at the BS. Furthermore, based on the CS channel measurements, we develop techniques to update and learn such channel AoD statistics at the BS. We use the upper confidence bound (UCB) algorithm to learn the AoD statistics and the CS matrix. Numerical results show that the CS matrix in the proposed framework provides faster beam alignment than standard CS matrix designs. Simulation results indicate that the proposed beam training technique can reduce overhead by 80% compared to exhaustive beam search, and 70% compared to standard CS solutions that do not exploit any AoD statistics.

## I. INTRODUCTION

With the availability of a large amount of bandwidth, mmWave communication can support massive sensor data sharing in vehicular networks [1]–[3]. Due to the large antenna arrays deployed at mmWave and the high mobility in vehicular settings, fast and efficient beam alignment strategies have to be designed [4]. For example, exhaustive search-based beam alignment where a radio scans through all the beams from a pre-defined codebook can result in substantial training overhead. The exhaustive search-based solution may not be suitable for mmWave vehicular communication due to a large number of candidates in typical beam codebooks. The mobile nature of terminals in a vehicular setting can impose a strict timing constraint for the beam alignment problem. As a result, exhaustive search-based beam alignment may not meet requirements like low latency and high reliability expected in mmWave vehicular networks [5], [6].

Fast beam alignment in mmWave communication can be achieved through compressed sensing (CS) or side information-aided beam search. CS-based beam alignment solutions exploit sparsity of mmWave channels and require

less training overhead than exhaustive search [7], [8]. The performance of CS-based solutions depends on the choice of the CS matrix used to obtain channel measurements. Side information, *ubiquitous* in vehicular communication, can be exploited for efficient mmWave channel estimation or beam alignment [8]–[13]. Different types of side information can be obtained from sensors mounted on vehicles, including radar, LiDAR and GPS [14]–[18]. Side information can be shared among vehicles by different wireless protocols such as dedicated short-range communication (DSRC), LTE or mmWave communication [1]. Situational awareness of vehicles can also be leveraged to extract environment information and further narrow down the beam search space and reduce beam training overhead [11], [12], [19]–[21]. One limitation of the side information-aided beam selection solutions is that they rely on fully-connected vehicles to share sensor data and enable sensor fusion. Second, none of the approaches leverage the spatial structures associated with vehicular channels for beam selection.

Vehicles channels are temporally correlated, which makes online learning a potential candidate for link configuration in vehicular settings [22]–[24]. Using online learning, a radio can dynamically adapt to the new channel statistics. In [25], beam selection was performed using vehicle arrival directions. In [21] and [26], receiver location was used for online beam selection in mmWave vehicular communication. Such online learning-based beam alignment solutions, however, do not exploit the sparsity of mmWave channels. Algorithms that integrate online learning-based solutions with sparsity-aware solutions may further reduce the training overhead.

In this paper, we assume a vehicular communication setting where a BS serves vehicles in its coverage area. We propose an online beam alignment framework that leverages channel sparsity and the angle domain statistics of the channel. Using our framework, the BS learns the AoD distribution of the paths corresponding to the vehicles in its coverage area. The AoD distribution learned at the BS is further used to optimize the CS matrix. It is important to note that the proposed approach only uses CS-based channel measurements acquired at mmWave, which are used for beam alignment. As a result, our framework does not require fully-connected vehicles, frequent sensor or out-of-band information sharing [8], [21], [26]. We assume that AoD distribution learning and CS-based computations are implemented at the centralized infrastructure, i.e., the BS. The main contributions of our work can be summarized as follows.

- We propose a novel solution for mmWave beam alignment leveraging an online learning-based CS matrix design. We develop a 2D-convolutional CS (2D-CCS)

Yuyang Wang, Nitin Jonathan Myers are with the Department of Electrical and Computer Engineering, the University of Texas at Austin, Austin, TX, 78712 USA, email: {yuywang, nitinmyers}@utexas.edu. Nuria González-Prelcic, and Robert W. Heath Jr. are with the Department of Electrical and Computer Engineering, North Carolina State University, 890 Oval Dr, Raleigh, NC 27606 USA, email: {ngprelcic, rwheathjr}@ncsu.edu. This material is based upon work supported in part by the National Science Foundation under Grant No. ECCS-1711702, and by a Qualcomm Faculty Award.

technique that optimize the CS matrix based on the AoD distribution learned at the BS. According to the CS matrix design, the BS applies beam training vectors for the users to acquire compressive channel measurements. Subsequently, the BS updates the AoD statistics based on the channel feedback received from the users. The optimized CS matrix can be obtained when the estimation of AoD distribution converges.

- We show that online CS matrix learning using our framework is analogous to a MAB problem [27]. We use the UCB algorithm for online CS matrix learning and investigate different exploration-exploitation methods to estimate the AoD distribution [28]. Furthermore, we propose solutions to calculate the CS matrix using a variation of the estimated AoD distribution with a reduced number of measurements. We apply different exploration terms to UCB based on the levels of *confidence* in the AoD distribution learning.
- We validate our proposed solution with comprehensive simulations. We demonstrate superior performance of the proposed solution over *exhaustive beam search* and *conventional CS* using the exact AoD prior. We show that the proposed online learning-based AoD distribution estimation technique converges to the ground truth using a *statistical distance* metric. Last, we evaluate the effectiveness of the proposed exploration-exploitation methods.

We would like to highlight that our online learning-based solution is different from the common adaptive CS setting in which the rows of a CS matrix are modified during signal acquisition [29]. Such an approach requires continuous feedback and may result in a substantial training overhead. In our framework, the CS matrix at the BS remains constant during the beam alignment process of any receiver, and is updated across different channel instances.

The rest of the paper is organized as follows. The problem is motivated in Section II, which demonstrates a specific structure in the AoD distribution at the BS that can be leveraged for data-driven CS in mmWave vehicular communication. The system model is explained in Section II-A. The channel model is provided in Section II-B. A new CS solution that is well-suited to the AoD prior is proposed in Section III. The online CS matrix learning problem is introduced in Section IV. The application of UCB for the online sensing matrix learning is shown in Section IV-A and Section IV-B. The exploration-exploitation tradeoff is explained in Section IV-C. Comprehensive numerical results are demonstrated in Section V. The final conclusions are drawn in Section VI.

**Notation:**  $\mathbf{A}$  is a matrix,  $\mathbf{a}$  is a column vector and  $a, A$  denote scalars. Using this notation  $\mathbf{A}^T$ ,  $\bar{\mathbf{A}}$  and  $\mathbf{A}^*$  represent the transpose, conjugate and conjugate transpose of  $\mathbf{A}$ . The scalar  $a_m$  or  $\mathbf{a}[m]$  denotes the  $m^{\text{th}}$  element of  $\mathbf{a}$ . The  $\ell_2$  norm of  $\mathbf{a}$  is denoted by  $\|\mathbf{a}\|_2$ . The scalar  $\mathbf{A}(k, \ell)$  denotes the entry of  $\mathbf{A}$  in the  $k^{\text{th}}$  row and  $\ell^{\text{th}}$  column. The matrix  $|\mathbf{A}|$  contains the element-wise magnitude of  $\mathbf{A}$ , i.e.,  $|\mathbf{A}|_{k,\ell} = |\mathbf{A}_{k,\ell}|$ . The  $\ell_1$  norm of  $\mathbf{A}$  is denoted by  $\|\mathbf{A}\|_1$ . The Frobenius norm is denoted by  $\|\mathbf{A}\|_F$ . The inner product of two matrices  $\mathbf{A}$  and  $\mathbf{B}$  is defined as  $\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{k,\ell} \mathbf{A}(k, \ell) \bar{\mathbf{B}}(k, \ell)$ . The

symbols  $\odot$  and  $\otimes$  are used for the Hadamard product and 2D circular convolution [30]. The set  $[N]$  denotes the set of integers  $\{0, 1, 2, \dots, N-1\}$ . The matrix  $\mathbf{U}_N \in \mathbb{C}^{N \times N}$  denotes the unitary Discrete Fourier Transform (DFT) matrix. Denote  $\omega_N = e^{-2j\pi/N}$ , we define the DFT matrix as  $\mathbf{U}_N(j, k) = \frac{\omega_N^{jk}}{\sqrt{N}}$ ,  $j, k \in [N]$ . Cross product of two sets is denoted as  $\times$ .

## II. MOTIVATION AND DATASET ESTABLISHMENT

In this section, we first explain the motivation of the work. We then describe our ray tracing simulation setup developed to collect vehicular channel data. Finally, we describe how the channels are computed from the ray tracing data to generate the channel dataset.

Channel statistics in vehicular communication follows regular patterns that are dependent on the site-specific environments [31]–[33]. We consider a scenario where a BS is deployed at road side to provide data services to passing-by vehicles that are equipped with mmWave communication devices. In most cellular scenarios, users can be located anywhere in the *whole* coverage area. For vehicular communication, however, vehicles travel on predetermined lanes that occupy only a *small portion* of the entire coverage area. In this case, the distribution of vehicle locations follows a statistical pattern that is related to the road layouts. Such statistical pattern of vehicle location distribution provides some informative and useful prior on the AoD corresponding to BS-vehicle links. In this paper, we mainly consider the transmit-side beam alignment in mmWave vehicular communication. We can observe from the dataset that there is certain statistical pattern in the AoD distribution related to street layout, which is associated with the stationary roadside BS. For the receiver side, however, since vehicles are highly mobile and are distributed at different locations, it is difficult to model the angle prior at the receiver side without the availability of accurate receiver location. Furthermore, even with the availability of the receiver location, it is not easy to predict the angle-of-arrival (AoA) in the channel. This is because the AoA also depends on the orientation of the receive array.

The availability of an informative AoD prior in mmWave vehicular communication settings can be helpful for efficient beam alignment or channel estimation algorithms design. For example, weighted sparse recovery techniques that exploit such a prior can be used instead of standard CS algorithms for mmWave link configuration [8]. Furthermore, the beam training vectors used to acquire channel measurements can be optimized to increase the probability of successful alignment. Determining the AoD prior in vehicular communication scenarios, however, can be challenging. A ray tracing-based approach to compute the AoD prior can be prohibitive in practical settings, as it may require accurate models for the surfaces of all the reflectors. The framework proposed in this paper learns the AoD prior online using compressed channel measurements, and can potentially adapt to new environments.

### A. Simulation setup

We use Wireless Insite, a commercial ray tracing simulator, to establish the channel dataset for evaluation. Ray tracing

simulation projects rays from the BS to the physical environment and calculates the channel path information, including path gain, angle-of-arrival (AoA) and AoD between the BS and the vehicles. We consider a street layout with two straight lanes in an urban canyon.

In the simulation, the buildings are modeled by cuboids with a concrete exterior. The simulation includes two types of vehicles: high trucks and low cars (e.g., sedans). For simplicity, all the vehicles are modeled as cuboids with *metal* exteriors. In each simulation, the type of vehicle (truck and car) is determined by a Bernoulli random variable with a predefined probability. The distance between adjacent vehicles is modeled by an *Erlang distribution* to form a *sequence* of vehicles on both lanes. More details of the ray tracing setup can be found in [12]. Our framework can be applied in *any* realistic vehicular environment as it learns the underlying AoD distribution from the channel measurements. We assume that the BS is mounted on top of street-side lamp-posts, which are higher than buses and are dotted around in cities. Receivers are placed on the roof of the vehicles. A particular realization of this simulation setup is shown in Fig. 1.



Fig. 1. Illustration of the ray tracing setup in which cars and trucks are randomly dropped in the two lanes of the urban canyon. Receivers are mounted on the top center of the low vehicle. The BS is mounted on a street-side lamp-post. The figure illustrates the *top five* strongest paths of the channel for a certain receiver. Our channel model includes the effect of multiple reflections that occur at the buildings and the vehicles.

### B. Channel model

We assume that the BS is equipped with a uniform planar array (UPA) of size  $N_x \times N_y$ . For ease of notation, we assume that the UPA has the same number of antennas along the azimuth and elevation dimensions, i.e.,  $N_x = N$  and  $N_y = N$ . The derivation, however, is not subject to the square antenna constraint. We consider an analog beamforming architecture at the BS. In such an architecture, every antenna of the UPA is connected to a single radio frequency chain through a phase shifter. The use of phase shifters allows the BS to generate a variety of beams that can be used for initial access or data transmission.

Therefore, under the multiple-input single output (MISO) system assumption, the antenna domain channel can be represented as a vector  $\mathbf{h} \in \mathbb{C}^{N^2 \times 1}$ . The channel vector  $\mathbf{h}$  can be reshaped into a matrix  $\mathbf{H} \in \mathbb{C}^{N \times N}$ . The rows and the columns

of  $\mathbf{H}$  correspond to the elevation and the azimuth dimensions of the UPA. Note that the  $(r, c)^{\text{th}}$  entry of  $\mathbf{H}$  represents the channel coefficient between the  $(r, c)^{\text{th}}$  antenna at the transmitter and the receiver antenna. The matrix representation of the MISO channel allows us to better explain the ideas underlying our beam training design. We consider both LOS and NLOS channels in our simulations.

We assume that the channel can be modeled using  $L_p$  paths in the propagation environment. We denote the channel gain of the  $\ell$ -th path as  $\alpha_\ell$ , the phase as  $\beta_\ell$ , the azimuth AoD as  $\phi_\ell$ , and the elevation AoD as  $\theta_\ell$ . We assume a half-wavelength spaced UPA at the BS to define the Vandermonde vector  $\mathbf{a}_N(\Delta)$  as

$$\mathbf{a}_N(\Delta) = [1, e^{j\pi\Delta}, e^{2j\pi\Delta}, \dots, e^{(N-1)j\pi\Delta}]. \quad (1)$$

Assuming the channel is narrowband, the MISO channel between the BS and the receiver can be expressed as

$$\mathbf{H} = \sum_{\ell=1}^{L_p} \alpha_\ell e^{j\beta_\ell} \mathbf{a}_N(\cos \theta_\ell) \mathbf{a}_N(\sin \theta_\ell \cos \phi_\ell)^T. \quad (2)$$

It should be noted that we can extend our framework to the wideband case using the frame structure in [34]. The techniques proposed in this paper estimate the  $N \times N$  channel matrix in (2). This approximation is used to update the AoD distribution, which is further used for efficient beam alignment.

The BS applies different phase shift matrices to its antenna array in  $M \ll N^2$  successive training slots for channel estimation. The receiver measures the projections of the channel on each of the phase shift matrices used by the BS. Finally, the receiver feedbacks the vector of  $M$  channel measurements to the BS for beam alignment in the control channel. In the  $m$ -th measurement slot,  $1 \leq m \leq M$ , the phase shift matrix is represented as  $\mathbf{P}[m] \in \mathbb{C}^{N \times N}$ . The matrix  $\mathbf{P}[m]$  has a unit norm and satisfies the constant amplitude constraint. Defining  $\mathbf{1}$  as an all-one matrix of size  $N \times N$ , we have  $|\mathbf{P}| = \mathbf{1}/N$ . As a result,  $\|\mathbf{P}\|_F = 1$ . Denoting the additive white Gaussian noise as  $v[m] \sim \mathcal{N}_c(0, \sigma^2)$ , the received signal can be written as

$$y[m] = \langle \mathbf{H}, \mathbf{P}[m] \rangle + v[m]. \quad (3)$$

The number of channel measurements required to estimate the  $N \times N$  matrix  $\mathbf{H}$  can be reduced by using techniques that exploit available statistical pattern in  $\mathbf{H}$ .

The channel matrix is approximately sparse in the 2D-DFT dictionary at mmWave carrier frequencies [2] when a UPA is assumed. We define the beamspace channel  $\mathbf{X} \in \mathbb{C}^{N \times N}$  as the inverse 2D-DFT of  $\mathbf{H}$ , i.e.,

$$\mathbf{X} = \mathbf{U}_N^* \mathbf{H} \mathbf{U}_N. \quad (4)$$

The beamspace contains the *effective* channel coefficients seen when different 2D-DFT-based beams are used at the BS [35]. Thanks to the limited scattering characteristics at mmWave frequencies, the beamspace  $\mathbf{X}$  is approximately sparse. The support of the non-zero coefficients in  $\mathbf{X}$  is related to the AoD of the propagation rays in the channel. The sparsity of  $\mathbf{X}$  allows the use of CS-based algorithms for fast channel estimation or beam alignment. For the tractability of analysis,

we assume that  $\mathbf{X}$  is one-sparse. It should be noted, however, that the channels used in our simulation are not one-sparse. These channels are derived from the ray tracing simulator using (2). The one-sparse prior for  $\mathbf{X}$  is perhaps simplistic, but allows a tractable design of a CS matrix that is well-suited to the prior on  $\mathbf{X}$ . In particular, the prior is only updated based on the strongest channel component, i.e., the direction that is the strongest in the 2D-DFT beamspace. As a result, our approach cannot learn all the components of the channel and we leave the extension to the future work. The proposed CS matrix design, however, achieves good performance over standard constructions for practical channels. It is important to note that the beamspace channels in our simulations include leakage effects and are approximately sparse.

### III. CS DESIGN WITH PERFECT AOD PRIOR

In this section, we assume the availability of a known AoD prior and design an optimization technique that optimizes CS matrix based on the AoD prior using a 2D-CCS framework [34]. This optimized CS matrix can be used for more efficient channel measurements to improve the beam alignment performance. In Section IV, we will propose an online learning framework that learns such AoD prior and optimizes CS matrix accordingly.

#### A. 2D-convolutional CS: Motivation and Background

Channel measurements in 2D-CCS are the projections of the channel matrix onto 2D circulant shifted versions of a base matrix  $\mathbf{P}$  [34]. For  $M$  channel measurements with 2D-CCS, the BS applies  $M$  distinct 2D-circulant shifts of  $\mathbf{P}$  to its phased array. The set of the circulant shifts used to acquire the  $M$  measurements is denoted as  $\Omega$ , given by

$$\Omega = \{(r[1], c[1]), (r[2], c[2]), \dots, (r[M], c[M])\}. \quad (5)$$

The coordinates  $\{(r[m], c[m])\}_{m=1}^M$  are chosen at random without replacement from  $[N] \times [N]$ . Subsampling randomly is a standard approach for 2D-CCS [34]. Specific subsampling design will be beneficial for different problems. It is an interesting problem to optimize the subsampling coordinates as well, since there exists some underlying angular information in the dataset. The subsampling optimization problem, however, is not analytically straightforward and is beyond the scope of this paper. Therefore, we simply assume the coordinates are randomly selected. In the  $m$ -th training slot, the phase shift matrix  $\mathbf{P}[m]$  in 2D-CCS is obtained by circulant shifting the base matrix  $\mathbf{P}$  by  $r[m]$  rows and  $c[m]$  columns. The receiver acquires the projection of  $\mathbf{H}$  on  $\mathbf{P}[m]$  according to (3). The motivation to use 2D-CCS is two fold. First, the CS matrix in 2D-CCS can be parametrized just by the base matrix  $\mathbf{P}$  and the set  $\Omega$ . Such a parametrization results in fewer optimization variables compared to that in generic CS matrix optimization. Second, 2D-CCS can exploit the AoD prior very well to design a proper base matrix suited to site-specific layouts of the streets. As the magnitude of the 2D-DFT of a matrix is invariant to its 2D-circulant shifts [36], the 2D-DFT magnitude of all the phase shift matrices in 2D-CCS is exactly the same. The 2D-DFT magnitude of a phase shift

matrix represents its beam pattern sampled at discrete angular locations. A reasonable channel measurement strategy under a AoD prior is to use beam patterns that focus more power along the directions that are more likely to have the strongest components in the channels.

Now, we explain the mathematical idea underlying 2D-CCS. We define  $\mathbf{P}_{\text{FC}}$  as the flipped and conjugated version of  $\mathbf{P}$ , i.e.,  $\mathbf{P}_{\text{FC}}(k, \ell) = \overline{\mathbf{P}}(N-k, N-\ell) \quad \forall k, \ell$ . To demonstrate how 2D-CCS works, we obtain the 2D circular convolution of the channel and  $\mathbf{P}_{\text{FC}}$  as

$$\mathbf{G} = \mathbf{H} \circledast \mathbf{P}_{\text{FC}}. \quad (6)$$

The  $k, \ell$ -th entry  $\mathbf{G}(k, \ell)$  is the projection of the channel onto the phase shift matrix obtained by shifting  $\mathbf{P}$  by  $k$  rows and  $\ell$  columns. We use  $\mathcal{P}_{\Omega}(\cdot) : \mathbb{C}^{N \times N} \rightarrow \mathbb{C}^{|\Omega| \times 1}$  to represent the projection operator that returns the entries of a matrix at the locations in  $\Omega$  as a vector. We define the noise vector  $\mathbf{v} \in \mathbb{C}^{|\Omega| \times 1}$ . With this definition, the channel measurements in 2D-CCS can be expressed as

$$\mathbf{y} = \mathcal{P}_{\Omega}(\mathbf{H} \circledast \mathbf{P}_{\text{FC}}) + \mathbf{v}. \quad (7)$$

Therefore, it can be observed that the channel measurements in 2D-CCS are a subsampling of the convolution between  $\mathbf{H}$  and  $\mathbf{P}_{\text{FC}}$ .

We use the mask concept to indicate the 2D-DFT of the phase shift matrix. From (8), we conclude that if we look at the measurements in the 2D-DFT basis, it can simply be represented as the *Hadamard* product between the beamspace  $\mathbf{X}$  and the 2D-DFT of the phase shift matrix  $\mathbf{P}$ . Therefore, the measurements can be obtained by applying a *mask* to the original beamspace to be measured in the 2D-DFT domain. CS-based channel estimation attempts to recover the beamspace based on the measurements. We further transform the CS problem corresponding to (7) into a partial 2D-DFT CS problem [34]. The transformation will be used to simplify our CS matrix optimization problem. Based on the DFT property, the inverse 2D-DFT of  $\mathbf{G}$  is given by [36]

$$\mathbf{U}_N^* (\mathbf{H} \circledast \mathbf{P}_{\text{FC}}) \mathbf{U}_N^* = \underbrace{\mathbf{U}_N^* \mathbf{H} \mathbf{U}_N^*}_{\mathbf{X}} \odot \underbrace{N \mathbf{U}_N^* \mathbf{P}_{\text{FC}} \mathbf{U}_N^*}_{\mathbf{Z}}, \quad (8)$$

which is represented as the Hadamard product of two terms,  $\mathbf{X}$  and  $\mathbf{Z}$ . The left side of the Hadamard product in (8) is the channel beamspace. The right side of the Hadamard product is defined as the *mask*. As  $\|\mathbf{P}\|_{\text{F}} = 1$ , it can be observed that  $\|\mathbf{Z}\|_{\text{F}} = N$ . We name the Hadamard product of  $\mathbf{X}$  and  $\mathbf{Z}$  as the masked beamspace [34]. It can be observed from (8) that  $\mathbf{H} \circledast \mathbf{P}_{\text{FC}} = \mathbf{U}_N (\mathbf{X} \odot \mathbf{Z}) \mathbf{U}_N$ . We substitute the 2D-DFT representation of (8) in (7), which gives

$$\mathbf{y} = \mathcal{P}_{\Omega} (\mathbf{U}_N (\mathbf{X} \odot \mathbf{Z}) \mathbf{U}_N) + \mathbf{v}. \quad (9)$$

As the masked beamspace  $\mathbf{X} \odot \mathbf{Z}$  is sparse, it can be recovered from its subsampled 2D-DFT in (9) with standard partial 2D-DFT techniques. To guarantee successful recovery of the beamspace  $\mathbf{X}$ , however, the mask  $\mathbf{Z}$  should be non-zero at all locations. To satisfy such non-zero constraints, a *unimodular* mask was used in [34]. Denoting the estimated beamspace as  $\hat{\mathbf{X}}$ , it was shown that the use of a unimodular mask minimizes the channel reconstruction error  $\|\mathbf{X} - \hat{\mathbf{X}}\|_{\text{F}}$  [34].

Non-uniform masks can be more efficient to acquire compressive channel measurements when the AoD prior is known at the BS. As the AoD prior in vehicular settings follows a certain statistical pattern according to the street layouts, designing a mask  $\mathbf{Z}$  that is tailored to the prior can result in better beam alignment. In this paper, we define the AoD prior as a probability distribution of the best beam index within the 2D-DFT beamspace. A good mask is one with a larger magnitude along the directions that are more likely to be optimal. As the squared Frobenius norm of the mask is constrained to be  $N^2$ , the mask design problem is analogous to a power allocation problem. In Section III-B, we investigate the mask design problem for a given AoD prior.

### B. Mask design with AoD prior

CS matrices in 2D-CCS are parameterized by  $\mathbf{P}$ , or the mask  $\mathbf{Z}$ , and the subsampling set  $\Omega$ . In this paper, we focus on the optimization over the magnitude of the mask and leave optimization of the subsampling set  $\Omega$  to future work. To this end, we assume  $\Omega = [N] \times [N]$ . Under such full sampling assumption, (9) can be rewritten as

$$\mathbf{y} = \text{vec}(\mathbf{U}_N(\mathbf{X} \odot \mathbf{Z})\mathbf{U}_N) + \mathbf{v}. \quad (10)$$

We define  $\tilde{\mathbf{y}}$  as an  $N^2 \times 1$  vector that is obtained by inverting the 2D-DFT in (10), i.e.,  $\tilde{\mathbf{y}} = (\mathbf{U}_N^* \otimes \mathbf{U}_N^*)\mathbf{y}$ . We define the transformed noise vector as  $\tilde{\mathbf{v}} = (\mathbf{U}_N^* \otimes \mathbf{U}_N^*)\mathbf{v}$  to write

$$\tilde{\mathbf{y}} = \text{vec}(\mathbf{X} \odot \mathbf{Z}) + \tilde{\mathbf{v}} \quad (11)$$

$$= \underbrace{\text{vec}(\mathbf{X})}_{\mathbf{x}} \odot \underbrace{\text{vec}(\mathbf{Z})}_{\mathbf{z}} + \tilde{\mathbf{v}}. \quad (12)$$

Due to the unitary nature of the 2D-DFT,  $\tilde{\mathbf{v}}$  has the same statistics as  $\mathbf{v}$ , i.e.,  $\tilde{\mathbf{v}} \sim \mathcal{N}_c(0, \sigma^2 \mathbf{I})$ . We assume that the one-sparse beamspace vector  $\mathbf{x}$  is an instance of a random variable  $\mathbf{x}$  that is supported on a discrete event set  $\mathcal{F}$ . Under the one-sparse prior assumption,  $\mathcal{F}$  contains the canonical basis vectors  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{N^2} \in \mathbb{C}^{N^2 \times 1}$ . The probability that  $\mathbf{x} = \mathbf{e}_k$ , i.e., the  $k^{\text{th}}$  beamspace component is 1, is  $p_k$ . The AoD prior is defined by the probabilities  $\{p_k\}_{k=1}^{N^2}$  that sum to one. The event that  $\mathbf{x} = \mathbf{e}_k$  corresponds to the case in which the  $k^{\text{th}}$  beam direction is optimal in the 2D-DFT beamspace.

Now, we derive the probability of successful beam alignment for a mask  $\mathbf{z}$ . For such a computation, we first derive the success probability conditioned on  $\mathbf{x} = \mathbf{e}_k$ . When the  $k^{\text{th}}$  beamspace direction is optimal, the entries of  $\tilde{\mathbf{y}}$  can be expressed as

$$\tilde{y}_k = 1 \cdot z_k + \tilde{v}_k = z_k + \tilde{v}_k, \quad (13)$$

$$\tilde{y}_j = 0 \cdot z_j + \tilde{v}_j = \tilde{v}_j, \quad \forall j \neq k. \quad (14)$$

For  $\mathbf{x} = \mathbf{e}_k$ , beam alignment is successful when  $|\tilde{y}_k| > |\tilde{y}_j|$ ,  $\forall j \neq k$ , i.e.,

$$|z_k + \tilde{v}_k| > |\tilde{v}_j|, \quad \forall j \neq k. \quad (15)$$

Such an event occurs with a probability  $\prod_{j \neq k} \mathbb{P}(|\tilde{v}_j| < |z_k + \tilde{v}_k|)$ . Considering the  $N^2$  candidates for  $\mathbf{x}$ , the probability of successful beam alignment is

$$\mathbb{P}(\text{successful alignment}) \quad (16)$$

$$\stackrel{(a)}{=} \sum_{k=1}^{N^2} p_k \prod_{j \neq k} \mathbb{P}(|\tilde{v}_j| < |z_k + \tilde{v}_k|) \quad (17)$$

$$\stackrel{(b)}{=} \sum_{k=1}^{N^2} p_k (\mathbb{P}(|\tilde{v}| < |z_k + \tilde{v}_k|))^{N^2-1} \quad (18)$$

$$= \sum_{k=1}^{N^2} p_k \left( \mathbb{P} \left( \left| \frac{\tilde{v}}{z_k} \right| < \left| 1 + \frac{\tilde{v}_k}{z_k} \right| \right) \right)^{N^2-1}, \quad (19)$$

where (a) is derived based on the assumption that the events in  $\mathcal{F}$  are independent, and (b) follows from the IID noise assumption.

We express the probability of successful beam alignment as a function of the AoD prior, the mask, and the noise variance  $\sigma^2$ . Conditioned on  $\mathbf{x} = \mathbf{e}_k$ , the random variable  $\frac{\tilde{v}}{z_k} \sim \mathcal{N} \left( 0, \frac{\sigma^2}{|z_k|} \right)$ . We use Lemma 1 to simplify (19).

**Lemma 1.** For  $x$  and  $y$  as IID complex Gaussian with  $x, y \sim \mathcal{N}(0, \xi^2)$ , we have

$$\mathbb{P}(|1+x|^2 \geq |y|^2) = 1 - \frac{1}{2} \exp \left( -\frac{1}{2\xi^2} \right). \quad (20)$$

*Proof.* For convenience of notation, we first rewrite  $\mathbb{P}(|1+x|^2 \geq |y|^2)$  as

$$\mathbb{P}(|1+x|^2 \geq |y|^2) = \mathbb{P} \left( \left| \frac{\sqrt{2}(1+x)}{\xi} \right|^2 \geq \left| \frac{\sqrt{2}y}{\xi} \right|^2 \right) \quad (21)$$

Since both  $x$  and  $y$  are complex Gaussian, the distribution of  $x_c^2 = \left| \frac{\sqrt{2}y}{\xi} \right|^2$  is a chi-square distribution with degree of 2.

The random variable  $x_{\text{nc}}^2 = \left| \frac{\sqrt{2}(1+x)}{\xi} \right|^2$  follows a non-centered chi-square distribution of degree 2 with  $\lambda = 2\xi^2$  [37]. The probability in (20) can be derived as

$$\begin{aligned} \mathbb{P}(|1+x|^2 \geq |y|^2) &= \mathbb{P}(x_c^2 \leq x_{\text{nc}}^2) = 1 - \mathbb{P}(x_c^2 \geq x_{\text{nc}}^2) \\ &= \mathbb{E}_{x_{\text{nc}}^2} \left[ 1 - \exp \left( -\frac{1}{2} x_{\text{nc}}^2 \right) \right] \stackrel{(a)}{=} 1 - \frac{1}{2} \exp \left( -\frac{1}{2\xi^2} \right), \end{aligned} \quad (22)$$

where (a) is derived based on the moment-generating function of non-centered chi-square distribution with degree 2 and  $\lambda = 2\xi^2$  [37].  $\square$

We use Lemma 1, to rewrite (19) as

$$\mathbb{P}(\text{successful alignment}) \quad (23)$$

$$= \sum_{k=1}^{N^2} p_k \left( 1 - \frac{1}{2} \exp \left( -\frac{|z_k|^2}{2\sigma^2} \right) \right)^{N^2-1}. \quad (24)$$

The goal of this section is to determine the mask  $\mathbf{z}$  that maximizes the probability in (23).

We now derive a convex optimization problem that maximizes an approximation of the beam alignment probability. As  $\|\mathbf{P}\|_F = 1$ , the mask must satisfy the norm constraint  $\|\mathbf{z}\|_2 = N$ . The mask optimization problem is then

$$\begin{aligned} \max \quad & \sum_{k=1}^{N^2} p_k \left( 1 - \frac{1}{2} \exp \left( -\frac{|z_k|^2}{2\sigma^2} \right) \right)^{N^2-1} \\ \text{s.t.} \quad & \sum_{k=1}^{N^2} |z_k|^2 = N^2. \end{aligned} \quad (25)$$

Unfortunately, the optimization problem in (25) is non-convex. Based on Jensen's inequality, the log-equivalent form of the objective in (25) can be lower-bounded by

$$\log \left( \sum_{k=1}^{N^2} p_k \left( 1 - \frac{1}{2} \exp \left( -\frac{|z_k|^2}{2\sigma^2} \right) \right)^{N^2-1} \right) \quad (26)$$

$$\geq \sum_{k=1}^{N^2} p_k \log \left( \left( 1 - \frac{1}{2} \exp \left( -\frac{|z_k|^2}{2\sigma^2} \right) \right)^{N^2-1} \right) \quad (27)$$

$$= (N^2 - 1) \sum_{i=1}^{N^2-1} p_k \log \left( 1 - \frac{1}{2} \exp \left( -\frac{|z_k|^2}{2\sigma^2} \right) \right). \quad (28)$$

Fortunately, the lower bound in (28) is convex in  $(|z_1|^2, |z_2|^2, \dots, |z_{N^2}|^2)^T$ , i.e.,  $|\mathbf{z}|^2$ . The constraint in (25) is also convex in the squared magnitude vector  $|\mathbf{z}|^2$ . The optimal mask  $|\mathbf{Z}|_{\text{opt}}$  that maximizes the lower bound on the probability of successful beam alignment can be obtained by the following optimization problem

$$\begin{aligned} \max \quad & \sum_{k=1}^{N^2-1} p_k \log \left( 1 - \frac{1}{2} \exp \left( -\frac{|z_k|^2}{2\sigma^2} \right) \right) \\ \text{s.t.} \quad & \sum_{k=1}^{N^2} |z_k|^2 = 1. \end{aligned} \quad (29)$$

The problem in (29) can be solved using standard convex optimization by considering  $\{|z_k|^2\}_{k=1}^{N^2}$  as the variables. We can observe that (29) shares similar structure to the classical *waterfilling* problem, by approximating the exponential function in (29) to its linear approximation. In this paper, we do not make such an approximation. Instead, we use the formulation in (29) to find the mask magnitude that is optimized for an AoD prior. Another way to interpret the optimization problem is optimizing the power allocated to measuring different beam directions. If no AoD prior is available to the BS, the intuitive way to design the mask is to make it uniform in all beam directions. In site-specific scenarios with AoD prior available, a smarter way is to design a mask with different magnitudes along different beam directions corresponding to its AoD distribution. Also, we can treat the mask as the beam pattern to measure the channel due to the Hadamard product property. The BS should design a nonuniform beam pattern to *probe* the beamspace based on the AoD prior. Hence, we explain it by allocating power to measure the beamspace. We use the CVXPY package in Python [38] to solve the convex optimization problem in (29).

### C. Phase recovery from the optimized mask magnitude

From (8), it can be observed that the phase shift matrix  $\mathbf{P}$  is the inverse 2D-DFT of the mask  $\mathbf{Z}$ . The optimization problem in (29), however, only provides the magnitude of the mask. It is important to note that the phase of the mask cannot be arbitrary. This is because the base phase shift matrix  $\mathbf{P}$ , i.e., the 2D-DFT of  $\mathbf{Z}$ , must satisfy the unit-norm constraint of  $|\mathbf{P}| = 1/N$ . The problem of finding a phase of a matrix such that its inverse 2D-DFT has a known magnitude has been investigated in optics [39]. The Gerchberg-Saxton

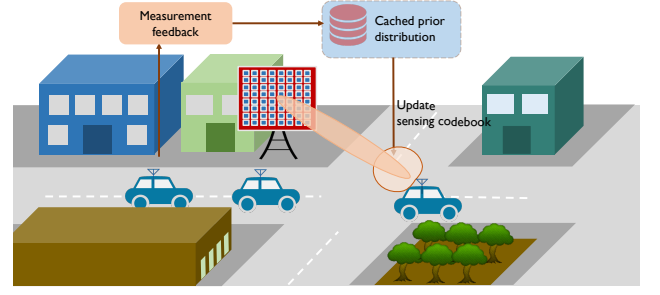


Fig. 2. An illustration of the online sensing matrix learning. The BS keeps updating and caching an AoD distribution corresponding to the vehicles in its coverage area. Given the available AoD prior, the BS designs its CS matrix for 2D-CCS based on the procedure in Section III and compressively measures the channel. Based on the channel measurement feedback, the BS estimates the direction with the strongest component in the 2D-DFT beamspace and updates the AoD prior accordingly. The BS keeps measuring the channels with a CS matrix optimized based on the available AoD prior and updating such AoD prior iteratively until the estimated AoD distribution converges.

algorithm in [39] iteratively retrieves the phase of a pair of signals related by Fourier transform, with the availability of the amplitudes of the two signals. In our problem, the base matrix  $\mathbf{P}$  and the mask  $\mathbf{Z}$  form a Fourier transform pair, with known amplitudes, i.e.,  $1/N$  and the optimized mask  $|\mathbf{Z}|_{\text{opt}}$ . The BS computes the phase of the mask  $\mathbf{Z}_{\text{opt}}$  and the corresponding base matrix  $\mathbf{P}_{\text{opt}}$  by the *Gerchberg-Saxton algorithm* [39]. The BS applies  $M \ll N^2$  random circulant shifts of the resultant base matrix  $\mathbf{P}_{\text{opt}}$  for the receiver to acquire 2D-CCS-based channel measurements.

## IV. ONLINE SENSING MATRIX LEARNING

In Section III, we explained how to design the sensing matrix in 2D-CCS with perfect AoD prior. Such prior can be obtained from a large number of channel realizations measured over a long time, which is similar to the procedure of *offline learning*. The establishment of such offline dataset, however, may not be easy. First, to identify the optimal beam index, a significant amount of overhead will be introduced due to the deployment of large antenna arrays and complicated channel statistics. Second, it is not clear how many data samples would be sufficient to estimate AoD distribution that is *accurate* enough for the design of the sensing matrix. Despite the changes in the traffic flow such as the traffic density and speed, we can tune the frequencies of CS matrix adaptations in different scenarios. In this paper, we show how online learning can be used to establish the angular domain database by learning the AoD distribution *on the fly*.

Online learning starts with the absence of any useful prior, i.e., the AoD distribution is uniform. The BS constructs the sensing matrix corresponding to the uniform AoD prior according to the procedure in Section III. The receiver feeds back the CS-based channel measurements to the BS. The BS then recovers the channel beamspace based on the channel feedback and updates the AoD distribution with the estimated optimal beam direction. Fig. 2 provides an illustration of the proposed online codebook learning solution. After the BS has successfully learned the underlying AoD prior, it applies the

optimized phase shift matrix for channel estimation and beam alignment using a small number of channel measurements. If a successful transmission link cannot be established using the proposed algorithm, the BS can simply reinitiate the beam training process based on the receiver feedback. The BS can also start the retransmission if the packet is not successfully received. We use a classical online learning solution in MAB, the UCB algorithm, to update the AoD prior.

In this paper, we use a stationary multi-armed bandits to formulate the online beam codebook learning framework. The BS learns the AoD prior and compressive beam codebooks through accumulated knowledge of channel prior information in a *site-specific* vehicular contexts, assuming a stationary AoD distribution. Although vehicular channels are highly dynamic, our solution focuses more on the relationship between the street topology (e.g. street layout) and channel angular prior, as shown in Fig. 4. Therefore, such stationary component of the angular prior has little dependence on the dynamics such as traffic condition, weather, etc. Instead, as long as the street topology and nearby buildings remain stationary, the channel prior will experience negligible changes. Furthermore, information such as traffic density is more related to the design of the problems such as scheduling or beam management. Acquisition of vehicular locations and other detailed information from sensor can significantly increase the network payload and also cannot guarantee sufficient sensor accuracy. This is also inconsistent with the motivation of the proposed AoD prior-aided beam alignment solution, where we would like to relax system requirements, as discussed in Section II. Therefore, we do not consider a typical reinforcement learning problem which deals with dynamic and adaptive vehicular contexts. We consider a stationary UCB model from multi-armed bandits.

#### A. Multi-armed bandits

A MAB is an online learning approach that learns the *policy* of actions and makes decisions from the *rewards* observed so far. The MAB problem can be explained as follows. There is a player with  $K$  independent machines, each with a reward distribution that is *initially unknown* to the player. The reward distributions of the  $K$  machines are assumed to remain constant at different time steps of the learning. The player plays an action  $a_t$  (tries one machine) from the given set  $\mathcal{A} = [K]$  and observes the associated reward  $R(t)$ . The action of the player at time  $t$  is only dependent on the past history of observed rewards  $H(t-1) = R(1), R(2), \dots, R(t-1)$ . A policy  $\pi$  is defined as the *mapping* from the history  $H(t-1)$  to the action  $a(t)$ . When the MAB algorithm becomes fully *confident* about its learned policy, it selects the machine that has the largest expected reward afterwards.

Online learning algorithms require a careful design of the *exploration* and *exploitation*. There exists a tradeoff between: 1) exploring the environment to find profitable actions, and 2) taking the empirically best action as often as possible. The tradeoff arises because only a limited number of machines are played at each time slot. The algorithm may repeat playing the optimal machine it observes so far from the rewards. Hence, the algorithm has to explore the environment to improve

its knowledge about the reward generating process. In UCB algorithms, such exploration-exploitation status is quantified by *upper confidence index* associated with each machine [27]. In [40], a family of policies was introduced where the index can be expressed as simple function of the sequence of the actions and rewards of each machine.

In UCB1 algorithm [40], for example, the upper confidence index is the sum of two terms: 1) the empirical mean of the rewards obtained from the sequence of observations, and 2) the *width* of the confidence bounds that reflects the uncertainty of the player's knowledge. In the basic setting, we assume a  $K$ -armed bandit problem [27]. We denote the time that arm  $k$  has been selected till time  $t-1$  as  $T_k(t-1)$ . The time slots in which arm  $k$  is played are  $\tau_1^k, \tau_2^k, \dots, \tau_{T_k(t-1)}^k$ . The empirical mean of the reward of  $k$ -th arm at time  $t$  is  $\hat{\mu}_k(t-1) = \frac{1}{T_k(t-1)} \sum_{j=1}^{T_k(t-1)} R(\tau_j^k)$ . For UCB1 algorithm, the upper confidence index  $\text{UCB}_k$  of the  $k$ -th machine at time  $t$  is defined as

$$\text{UCB}_k(t, \delta) = \begin{cases} \infty, & \text{if } T_k(t-1) = 0 \\ \underbrace{\hat{\mu}_k(t-1)}_{\text{Empirical mean}} + \underbrace{\sqrt{\frac{2 \log(1/\delta)}{T_k(t-1)}}}_{\text{Exploration term}}, & \text{if } T_k(t-1) > 0. \end{cases} \quad (30)$$

In each time step, the player plays the arm with the largest UCB value. The first term quantifies the knowledge the player accumulates so far as the empirical mean. The second term evaluates exploration as the square root of the inverse of the time the arm has been played. The algorithm suppresses the exploration for arms that have been tested many times, while *encourages* those arms that have been tried for only few times. The parameter  $\delta$  can be used to control the extent of the exploration [41]. When time  $t \rightarrow \infty$ ,  $T_k(t) \rightarrow \infty$ ,  $\forall k$  and the exploration term in (30) becomes zero. The UCB value becomes the empirical mean and converges to the expected rewards of the machines. The player simply chooses the arm with the largest expected reward at the end of the learning.

#### B. Online sensing matrix learning with UCB

The UCB algorithm fits the online beam alignment problem using a beam scanning-based approach, but its application to the online sensing matrix learning problem is not straightforward. For CS problems, the use of multi-armed bandits is not straightforward. In CS-based beam training approach, the beam used for measuring the channel is not directional. Instead, it tends to be wide and covers all beam directions or cover several beam directions with different weights based on the magnitude of the beam pattern along different beam directions. Therefore, it is no longer consistent with the concept of pulling one arm in multi-armed bandits. Furthermore, for classical multi-armed bandits problem or beam-scanning problem, pulling one arm or trying one beam generally returns the correct observation. For example, the player can check if pulling this arm brings rewards, or the BS can obtain the exact channel statistics along this beam direction. For CS-based solution, however, since the beam is wide, the measurements



will be a linear combination of the beam statistics along multiple beam directions. Leveraging sparsity can help to recover the beam RSRP along certain beam directions, without any guarantee of the channel estimation accuracy. This indicates that the observation of the channel statistics for *pulling* in CS-based solution is not accurate. Therefore, it is hard to characterize the notation of rewards in UCB.

Consider an example where the BS wants to test the  $k$ -th beam direction. Instead of obtaining feedback of the beamspace  $\mathbf{x}_k$  directly through beam search, the BS recovers an estimate of the whole beamspace  $\hat{\mathbf{x}}$  with CS. The *mismatch* between the estimated beamspace and the true beamspace, i.e.,  $\hat{\mathbf{x}} - \mathbf{x}$ , could pose great challenges to the exploration-exploitation design in online learning. The ultimate goal of the paper is to learn the sensing matrix that is best suited for a certain vehicular context. Since both the prior and the sensing matrix are stored and applied at the BS, the online codebook learning model introduces negligible signaling overhead and fully leverages the computation resources at the BS cloud.

### C. Exploration-exploitation in online sensing matrix learning

The online sensing matrix learning problem requires an elaborate design of the exploration-exploitation strategy. Without the availability of any useful prior, the online learning framework starts with a uniform mask. If the exploration term in (30) is not carefully designed, the BS may apply a uniform mask that does not fully leverage the empirical AoD distribution it learns so far. The use of a uniform mask can lead to poor channel estimation performance and slow convergence of the AoD prior to its real distribution. If the BS uses an AoD distribution that is estimated without sufficient exploration, it can end up using masks that may not have good magnitudes along unexplored directions that are likely to be optimal. The estimated AoD distribution could be more and more *biased* towards the beam directions that are estimated as optimal in the beginning of the learning. We propose three approaches to achieve the exploration-exploitation tradeoff. The first approach targets a smooth update of the AoD prior estimation and a *mild* exploration over different beams by adding an exploration term to the estimated AoD prior. The second approach, instead of imposing an exploration term to AoD estimation, requires the BS to directly apply such exploration term over the mask. Finally, we demonstrate that by changing the number of CS measurements over time, the proposed approach can achieve a more accurate AoD distribution estimation, especially in the beginning of online sensing matrix learning.

1) *UCB with exploration*: A good design of an exploration-exploitation tradeoff should guarantee a *smooth* evolvement of UCB values through the learning process. The learning starts with a uniform prior, but gradually learns more about the channel and converges to a *less uniform* AoD distribution that can be used for calculating a better mask. The 2D-CCS solution, however, cannot handle a *peaky* yet wrong prior. The exploitation term in UCB of (30), which is the empirical mean of the AoD distribution, is susceptible to AoD prior estimation inaccuracy due to an insufficient number of channel

realizations in the beginning. The time step  $t$  in the online learning is equivalent to the *index* of the sequential BS-vehicle channel realizations. Since the BS calculates the optimal beam direction based on CS measurements, the reward at time  $t$  is a one-hot encoded vector  $\mathbf{r}_t$ . With the estimated optimal beam direction at time  $t$  denoted as  $\ell(t)$ , we have  $\mathbf{r}_t[\ell(t)] = 1$  and  $\mathbf{r}_t[k] = 0, \forall k \neq \ell(t), k \in [N^2]$ . Under successful CS-based recovery, the BS compressively measuring the channel using 2D-CCS is treated as equivalent to the case where the BS measures all beam directions *exhaustively* and returns the optimal beam direction. Therefore, the BS forms a *superarm* that plays all machines each time, i.e.,  $T_k(t+1) = T_k(t) + 1, \forall i$ . With the existence of an identical exploration term for different beam directions in (30) that decreases over time, the mask is less *sensitive* to the change of AoD distribution estimation. Hence, the BS can avoid a peaky mask that arises from the inaccurate AoD estimation in the beginning. The slow decrease of the UCB *uniformity* guarantees the smoothness of the AoD prior at certain locations during initial exploration. The proposed algorithm is explained in Algorithm 1.

---

#### Algorithm 1 UCB for online sensing matrix learning.

---

**Input:** Number of measurements per time slot  $M$ , a constant  $\delta$ , number of plays  $T$ .

**Output:** Estimated AoD prior  $\hat{\mathbf{p}} = [\hat{p}_1, \hat{p}_2, \dots, \hat{p}_{N^2}]$ .

*Initialization:* Initialize a uniform AoD prior

$\mathbf{p} = [p_1, p_2, \dots, p_{N^2}]$ , where  $p_i = \frac{1}{N^2}, i = 1, 2, \dots, N^2$ .

A sufficiently large value is set as  $N_\infty = 10000$ .

Initialize  $T_i(1) = 0, \hat{\mu}_i(t) = 0, \forall i$ .

- 1: **for**  $t : \text{mod}(t, T_{\text{update}}) == 0$  **do**
  - 2:   Compute    UCB    values    UCB <sub>$i$</sub> ( $t$ )    ←  

$$\begin{cases} N_\infty, & \text{if } T_i(t-1) = 0 \\ \hat{\mu}_i(t-1) + \sqrt{\frac{2 \log(1/\delta)}{T_i(t-1)}}, & \text{if } T_i(t-1) > 0. \end{cases}$$
  - 3:   Update prior  $p_i = \frac{\text{UCB}_i(t)}{\sum_{j=1}^{N^2} \text{UCB}_j(t)}, \forall i$ , and calculate the sensing matrix based on Section III-A - III-C.
  - 4:   Perform 2D-CCS and return the estimated optimal beam direction index  $s$ .
  - 5:   Update reward and its empirical mean  
 $\mathbf{r}_t[i] = 0, \hat{\mu}_i(t+1) = \hat{\mu}_i(t) \forall i \neq s,$   
 $\mathbf{r}_t[s] = 1, \hat{\mu}_s(t+1) = \frac{\hat{\mu}_s(t)T_s(t)+1}{T_s(t)+1}.$
  - 6:   Update time of being played  $T_i(t+1) = T_i(t) + 1, \forall i$ .
  - 7: **end for**
  - 8:  $\hat{\mathbf{p}} = \mathbf{p}$
  - 9: **return** AoD distribution estimate  $\hat{\mathbf{p}}$ .
- 

2) *Mask regularization*: The second approach applies exploration directly on the mask  $\mathbf{Z}$ . In the first approach, we add an *exploration* term to the estimated AoD distribution in (30) to minimize the error in the AoD prior estimate that can result in a biased mask. In 2D-CCS, the magnitude of the mask represents the power allocated for sensing along different beam directions that are defined by the 2D-DFT dictionary. We propose to directly apply *regularization* over the mask. The BS calculates a mask directly based on the empirical AoD distribution, added by an extra exploration term in (30) as in Algorithm 1. The algorithm is provided in Algorithm 2. It can be observed that Algorithm 1 and



2 share very similar inputs, outputs, and procedures. The difference is where the exploration (regularization) term is imposed on. The regularization is inexplicitly applied for the AoD estimation of UCB in Section IV-C1, while it is directly incorporated in the mask for probability estimation. It should be noted from Algorithm 1 and Algorithm 2 that the BS does not update the optimized phase shift matrix for each time slot, since it could be expensive to solve the convex optimization in (29). We assume that the phase shift matrix is updated every  $T_{\text{update}} = 20$  channels. Such phase shift matrix optimization interval can achieve a good tradeoff between the system complexity and the algorithm efficiency.

3) *Reducing the number of CS measurements with online learning:* For the exploration-exploitation approaches proposed in Section IV-C2 and IV-C1, we assume a fixed number of CS measurements applied for each channel realization. Further varying the number of CS measurements is another way to improve AoD estimation accuracy and achieve a tradeoff between exploitation and exploration, without sacrificing too much in terms of overhead. When a larger number of channel measurements are acquired by the receivers, CS can reveal more accurate information in the beamspace. In the beginning of the online AoD distribution estimation, the BS must use a larger number of CS-based measurements to guarantee successful beamspace recovery. Successful CS-based beam alignment ensures that the BS obtains a more accurate AoD prior. With a better mask, fewer measurements are sufficient to recover the correct the optimal beam direction. Hence, we propose to apply Algorithm 1 and 2 while adapting the number of CS measurements in different *phases* of online learning. We consider the number of measurements as a function of time step  $t$  as  $M(t)$ .

The BS uses a higher number of CS measurements at the start of the learning, and decreases  $M(t)$  linearly with time. Specifically,  $\Delta_t$  defines the time interval after which the CS measurements  $M(t)$  is decreased by  $\Delta_M$ . It should be noted that the number of CS measurements cannot be too small so as to guarantee the accuracy of beam alignment. The initial number of CS measurements is represented as  $M_0$ , and the lower bound on the number of CS measurements is  $M_{\min}$ . Hence,  $M(t)$  changes over time and can be given by

$$M(t) = \max \left\{ M_0 - \left\lfloor \frac{t}{\Delta_t} \right\rfloor \Delta_M, M_{\min} \right\}. \quad (31)$$

Linear decrease of  $M(t)$  is a simple way of adjusting measurements temporally. For this approach, multiple hyperparameters, such as  $\Delta_t$ ,  $\Delta_M$ ,  $M_0$ , need to be tested and evaluated. In scenarios where there are large changes in the environment dynamics,  $M_{\min}$  measurements might not be enough to achieve reasonable system performance. In this case, a smaller  $M_{\min}$  that suits the changes in the environment must be used. Furthermore, the received SNR or rate can also be used as the reward instead of a unit reward in Algorithm 1. With these changes, the proposed algorithm starts with a uniform prior when it observes a poor reward in dynamic settings. Investigating a reasonable choice for  $M_{\min}$  for a vehicular scenario with a dynamic prior is an interesting direction for future work.

---

**Algorithm 2** Probability estimation with mask regularization

---

**Input:** Number of measurements per time slot  $M$ , a constant  $\delta$ , number of plays  $T$ .

**Output:** Estimated AoD prior  $\hat{\mathbf{p}} = [\hat{p}_1, \hat{p}_2, \dots, \hat{p}_{N^2}]$ .

*Initialization:* Initialize a uniform AoD prior  $\mathbf{p} = [p_1, p_2, \dots, p_{N^2}]$ , where  $p_i = \frac{1}{N^2}$ ,  $i = 1, 2, \dots, N^2$ .

A sufficiently large value  $N_{\infty} = 10000$ .

Initialize  $T_i(1) = 0$ ,  $\hat{\mu}_i(t) = 0$ ,  $\forall i$ .

- 1: **for**  $t : \text{mod}(t, T_{\text{update}}) == 0$  **do**
  - 2:   Update the estimated AoD distribution  

$$p_i \leftarrow \frac{\hat{\mu}_i(t)}{\sum_{j=1}^{N^2} \hat{\mu}_j(t)}.$$
  - 3:   Calculate the mask  $\mathbf{z}$  based on  $p_1, p_2, \dots, p_{N^2}$  based on Section III-A - III-C.
  - 4:   Regularize the mask amplitude with  

$$|\mathbf{z}_i| \leftarrow \begin{cases} |\mathbf{z}_i| + N_{\infty}, & \text{if } T_i(t-1) = 0 \\ |\mathbf{z}_i| + \sqrt{\frac{2 \log(1/\delta)}{T_i(t-1)}}, & \text{if } T_i(t-1) > 0. \end{cases}$$
  - 5:   Normalize the mask amplitude  $|\mathbf{z}|$ . Apply Gerchberg-Saxton Algorithm as in Section III-C to obtain the phase of the mask.
  - 6:   Perform 2D-CCS with  $\mathbf{z}$  and return the estimated optimal beam direction index  $s$ .
  - 7:   Update reward and its empirical mean  

$$x_i(t) = 0, \hat{\mu}_i(t+1) = \hat{\mu}_i(t) \quad \forall i \neq s, x_s(t) = 1,$$

$$\hat{\mu}_s(t+1) = \frac{\hat{\mu}_s(t)T_s(t)+1}{T_s(t)+1}.$$
  - 8:   Update the AoD distribution as the empirical mean  

$$p_i = \hat{\mu}_i(t), \quad \forall i.$$
  - 9:   Update time of being played for arm  $i$ ,  

$$T_i(t+1) = T_i(t) + 1, \quad \forall i.$$
  - 10: **end for**
  - 11:  $\hat{\mathbf{p}} = \mathbf{p}$
  - 12: **return** AoD distribution estimate  $\hat{\mathbf{p}}$ .
- 

## V. NUMERICAL RESULTS

In this section, we evaluate the proposed online 2D-CCS matrix learning framework with comprehensive numerical results. First, we examine the baseline model with perfect AoD prior, i.e., offline learning. Second, we compare performance of the online sensing matrix learning using different exploration-exploitation techniques proposed in Section IV-C. Last, we demonstrate the convergence of the estimated AoD prior with the proposed online sensing matrix learning algorithms.

### A. Offline learning with perfect AoD prior

In Fig. 3, we evaluate the beam alignment performance of 2D-CCS using perfect AoD prior. The performance metric is the average beam reference signal received power (RSRP). We consider a  $32 \times 32$  UPA at the BS. In the simulation, the transmit power at the BS is adjusted so that the average received SNR for a quasi-omnidirectional transmission is 10 dB. Such an SNR was achieved by using a Golay code to acquire channel measurements [34]. We compare the performance among the proposed 2D-CCS solution, exhaustive beam search with 2D-DFT beam codebook, and the baseline model

of optimal beamforming with perfect channel state information (CSI).

The development in this paper assumed a narrowband channel model, but the most interesting applications of mmWave use large bandwidths, and consequently have wideband channels. As a result, we propose a frame structure that allows applying the proposed 2D-CCS-based solution in the wideband case. First, we assume that a complementary Golay pair is used in the time domain for training, which is used in IEEE 802.11ay for example. Exploiting the complementary property, the receiver can separate out the different wideband channel taps. There are different ways to apply our algorithm with multiple channel taps  $\{\tilde{\mathbf{H}}_\ell\}_{\ell=0,1,\dots,L-1}$ . Because the channel taps have contributions primarily from a few angular directions, we simply add up the different matrix components. A more sophisticated approach could operate on each tap separately or jointly, but we defer this to future work. Also, it should be noted that we evaluate the channel based on the realistic ray tracing outputs, without the one-sparse assumption adopted in Lemma 1.

Let  $\tilde{\mathbf{y}}[\ell]$  denote the measurement vector corresponding to  $\tilde{\mathbf{H}}[\ell]$ . Our CS-based online learning framework is applied over the measurement vector  $\mathbf{y} = \sum_{\ell=0}^{L-1} \tilde{\mathbf{y}}[\ell]$ . As the analog beamtraining vectors are frequency flat, CS over  $\mathbf{y}$  is expected to provide an estimate of  $\mathbf{H} = \sum_{\ell=0}^{L-1} \tilde{\mathbf{H}}[\ell]$ . Let  $\mathbf{P}_{\text{BF}}$  represent the beamforming matrix applied to the phased array at the BS. The beamforming gain achieved with  $\mathbf{P}_{\text{BF}}$  is defined as

$$\text{BFgain} = |\langle \mathbf{H}, \mathbf{P}_{\text{BF}} \rangle|^2. \quad (32)$$

For the perfect CSI case with optimal beamforming, the beamforming matrix at the phased array is  $\mathbf{P}_{\text{BF, ideal}} = e^{j\text{phase}(\mathbf{H})}/N$ . We use  $\hat{\mathbf{H}}$  to denote the channel estimate obtained with our online learning-based 2D-CCS algorithm. The orthogonal matching pursuit (OMP) algorithm [42] was used to solve the sparse recovery problem in our method. The beamforming matrix with the proposed approach is  $\mathbf{P}_{\text{BF, 2D-CCS}} = e^{j\text{phase}(\hat{\mathbf{H}})}/N$ . For exhaustive beam search, the BS searches through the  $N^2$  beams from a 2D-DFT beam codebook and identifies the optimal beam direction with the largest beam RSRP. For the beam sweeping among the partial beamspace, with the availability of AoD distribution *ordered* as  $\mathbf{p} = [p_{r_1}, p_{r_2}, \dots, p_{r_{N^2}} : p_{r_1} > p_{r_2} > \dots > p_{r_{N^2}}]$ , the BS measures the RSRP of the *top M* beams and selects the one with the largest RSRP correspondingly.

Fig. 3 shows huge performance improvement with nonuniform mask designed by AoD prior, compared to that with uniform mask in the 2D-CCS framework. With 20 measurements, 2D-CCS with nonuniform mask shows more than 6 dB improvement of average beam RSRP than using beam search with the same number of measurements, and 10 dB gain compared to 2D-CCS with a uniform mask. With only 35 measurements, 2D-CCS with AoD prior achieves comparable performance with exhaustive search that uses  $N^2 = 1024$  beams. The average beam power using 2D-CCS-based beam alignment with AoD prior is superior to exhaustive beam search when  $M > 40$ , and almost achieves the same performance as perfect CSI case. It should be noted that the beam alignment with 2D-CCS outperforms exhaustive beam

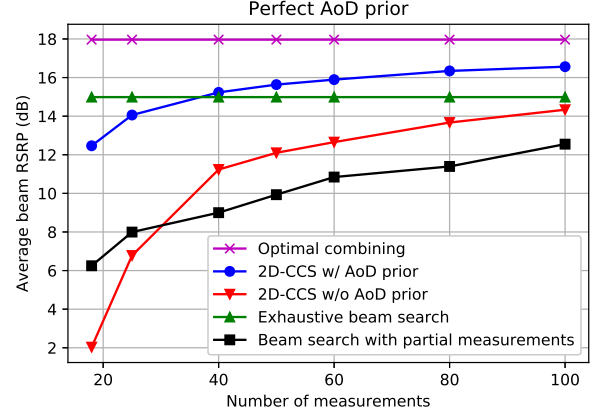


Fig. 3. A comparison of the average beam RSRP using different beam alignment approaches assuming perfect AoD prior. The proposed 2D-CCS with AoD prior outperforms that using uniform mask by a large margin. The average beam RSRP of 2D-CCS using AoD prior is close to that of optimal beamforming. Furthermore, 2D-CCS with AoD prior exhibits larger beamforming gain compared to exhaustive beam search, when the number of measurements grows large.

search because the BS applies conjugate beamforming using the estimated beamspace.

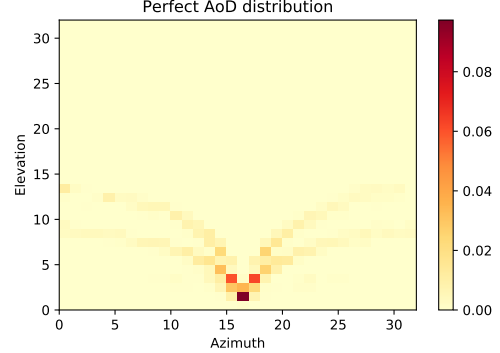


Fig. 4. AoD distribution in beamspace across all channel realizations.

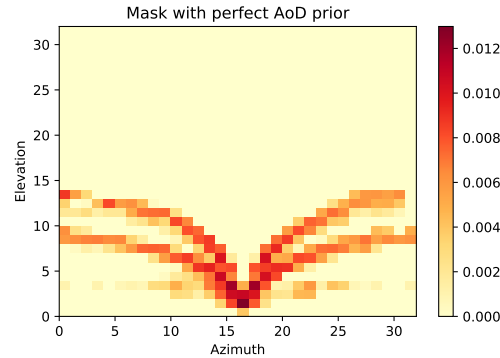


Fig. 5. Mask calculated using the optimization procedure in Section III for the AoD prior in Fig. 4

The real AoD distribution and the corresponding mask are plotted in Fig. 4 and Fig. 5. Two *continuous strips* can be observed in Fig. 4 and Fig. 5. These strips represent the two straight lanes in the simulated urban canyon. The observation

implies that structured AoD distribution can be extracted from site-specific street layouts and leveraged for CS matrix design. Furthermore, as shown in Fig. 5, the mask obtained by the optimization in (29) satisfies two desired properties: i) The BS is able to allocate more power to compressively sense the *dominant* beam directions; ii) The mask is also able to cover *less important* beam directions, e.g., the beamspace on the two sides of the strips along the azimuth direction in Fig. 4.

### B. Online sensing matrix learning

In this section, we evaluate and compare the effect of “*exploration*” on the online learning performance. We consider the case without any exploration, i.e., directly using the mask calculated by the estimated AoD distribution for sensing, and the three approaches provided in Section IV-C.

1) *No exploration*: Fig. 6 compares the mask learned without any exploration, using a fixed number of measurements. Without exploration, the procedure is identical to Algorithm 1 where  $\delta = \infty$  and the exploration term is zero. In this case, the sensing matrix is calculated directly based on the empirical AoD distribution without any regularization. As discussed in Section IV-C, no exploration leads to error propagation and highly biased estimation, which are reflected in the results in Fig. 6. Compared to the mask calculated from perfect AoD prior, the mask without exploration displays some *inconsistency* in the AoD distribution estimation, and fails to identify several important beam directions. The combined effect of measurement inaccuracy and error propagation due to biased AoD estimation gives rise to a mask that results in poor performance with CS.

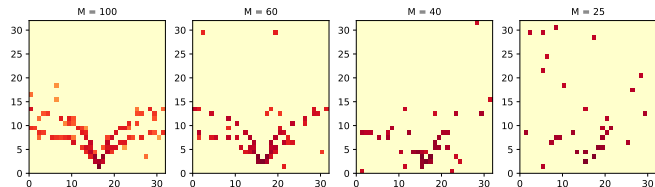


Fig. 6. A comparison of the final mask learned without any exploration, using different numbers of beam measurements  $M = 100, 60, 40, 25$ . Without exploration, the mask fails to capture the full channel statistics in the angular domain, especially when the number of CS measurements per channel realization is relatively small, e.g.,  $M = 25$  or  $M = 40$ .

2) *Exploration and regularization*: Fig. 7 and Fig. 8 illustrate the average beamforming loss with UCB in Section IV-C1 and mask regularization in Section IV-C2, compared to the *ideal mask* calculated from perfect AoD prior in Fig. 5 using different number of measurements. Specifically, we evaluate the performance of our approach for different values of  $\delta$ . Fig. 8 shows that the gap between the beamforming gain of the mask with perfect AoD prior and the one learned online using a fixed number of measurements is marginal, if the regularization parameter  $\delta$  is selected carefully. We observe that a low beamforming loss is achieved when  $\sqrt{2 \log(1/\delta)}$  is around 0.1. When  $\sqrt{2 \log(1/\delta)}$  is too large, e.g.,  $\sqrt{2 \log(1/\delta)} = 1$ , the mask changes too slowly to capture the information provided by AoD estimation. When  $\sqrt{2 \log(1/\delta)}$  becomes too small,

the learning could be too *aggressive* in the AoD distribution estimation, since it results in a poor mask for 2D-CCS. Similar results of UCB exploration can be observed in Fig. 7.

Table I provides a more explicit comparison among the beamforming loss of the mask with perfect AoD prior, uniform mask, the mask achieved with UCB exploration, and the mask with mask regularization, with respect to optimal beamforming using perfect CSI. The beamforming loss decreases with a larger number of CS measurements per channel realization. For  $M = 40, 50, 60, 80, 100$ , the optimal beamforming loss with respect to mask with perfect AoD prior is approximately 1 dB. When  $M = 25$ , the regularized mask has superior performance to that of the uniform mask. Also, using only 25 measurements throughout the whole learning process, the BS is still able to learn the AoD distribution and results in 2.19 dB beamforming loss compared to the mask calculated using perfect AoD prior. Similar observations can be obtained for different number of CS measurements.

We measure the AoD estimation accuracy based on statistical distance between the estimated AoD distribution and the real AoD distribution of the data samples [43]. We use a statistical distance measure - Hellinger distance [44]. For two discrete distributions  $P = (p_1, p_2, \dots, p_K)$  and  $Q = (q_1, q_2, \dots, q_K)$ , the Hellinger distance  $\mathcal{H}(P||Q)$  is defined as

$$\mathcal{H}(P||Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^K (\sqrt{p_i} - \sqrt{q_i})^2}. \quad (33)$$

The Hellinger distance is directly related to the Euclidean norm of the difference of the square root of the probability distribution probability mass functions (PMF).

Fig. 9 characterizes the temporal evolution of the AoD estimation with mask regularization using a fixed number of measurements. We compare the Hellinger distance between the empirical AoD distribution and the real AoD distribution, with mask regularization, and that with perfect CSI available at the BS. The real AoD prior is calculated across all 3000 channel realizations. The empirical AoD prior is calculated at different time steps, e.g., the AoD prior of the first 200, 300 samples, etc. For the perfect CSI case, we assume the BS can always recover the optimal beam direction. The perfect CSI curve provides a baseline *reference* to evaluate the AoD estimation accuracy *during* online learning. In Fig. 9, the Hellinger distance decreases sharply in the beginning. When time  $T < 500$ , the Hellinger distance with  $M = 100, 50$  closely matches that with perfect CSI. After time  $T > 1000$ , the Hellinger distance starts to change slowly. It is also shown that the Hellinger distances of  $M = 100$  and  $M = 50$  are similar. When  $M = 20$ , online learning cannot accurately estimate the exact optimal beam direction, which leads to a large Hellinger distance compared to the perfect CSI case. From the simulation results in Fig. 3 and Table I, we can see that the learned mask can well capture the underlying statistics in the AoD prior, and well resembles the perfect mask optimized based on the real AoD prior, despite the existence of the nonzero Hellinger distance. Therefore, the existence of such existent gap in the Hellinger distance will not impose

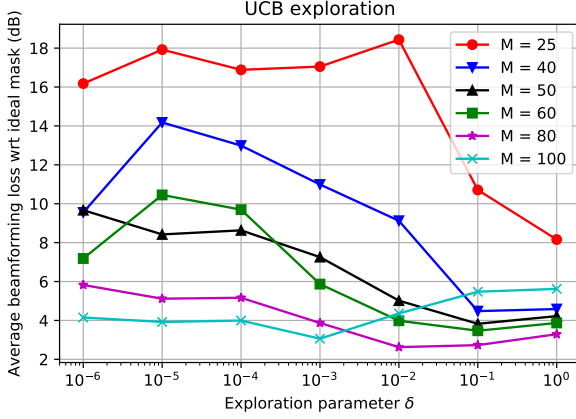


Fig. 7. Comparison of the average beamforming loss using UCB exploration in Section IV-C1 with respect to the performance achieved by applying mask in 2D-CCS with perfect AoD prior.

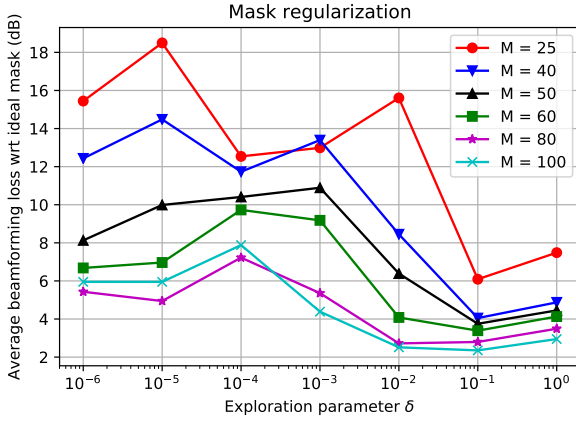


Fig. 8. Comparison of the average beamforming loss using mask regularization in Section IV-C2 with respect to the performance achieved by applying mask in 2D-CCS with perfect AoD prior.

large impacts on the system performance. Also, the Hellinger distance converges very fast to a certain value. The reason for this fast convergence is that channels collected in our dataset is based on ray tracing simulation in an urban canyon with two straight lanes. The conclusion implies that the proposed AoD prior-based sensing matrix design is *robust* to small AoD distribution estimation error, which is favorable to real field implementation that generally has more complicated street layouts and channel characteristics.

3) *Updating the number of CS measurements*: The online sensing matrix learning is built on the assumption that the CS measurements can roughly give accurate channel estimation, especially in the beginning of the learning. Starting with a large number of measurements when the BS has no prior is a good way to avoid error propagation in the AoD distribution estimation. After measuring more channel realizations, the BS becomes confident about the AoD distribution estimate. In such a case, decreasing the number of CS measurements can result in a reduced training overhead. With mask regularization, the BS further decreases the number of CS measurements in our online learning-based framework.

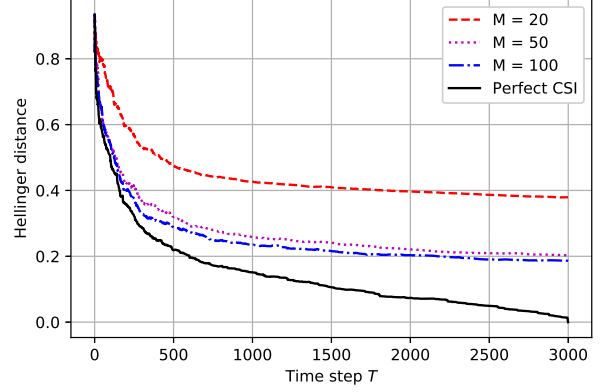


Fig. 9. Evaluation of the Hellinger distance defined in (33) between the real AoD distribution of in total of 3000 channel realizations, and the empirical AoD distribution learned at different time steps during online sensing matrix learning using mask regularization in Section IV-C2. The regularization parameter is set as  $\delta = 10$ , which is able to achieve a good exploration-exploitation tradeoff as shown in Fig. 7 and Fig. 8.

We present the average beamforming loss with a decreasing number of measurements in the 4 - 7 rows in Table I. We use  $M = M_{\min}$ , i.e., the *stabilized* number of measurements the BS uses for each channel realization. In our evaluation, we fix  $\Delta_M = 20$ . We update  $\Delta_t$  in (31) to modify how fast  $M$  decreases. The learning is shown to outperform mask regularization with  $M = M_{\min}$ . For  $M_{\min} = 25$ , an extra 1 dB gain compared to regularized mask can be achieved. No significant differences, however, are observed among the different initial measurement numbers  $M_0 = 200, 300$  and the update interval  $\Delta_t = 100, 200$ . The *worst-case* scenario with the *smallest* initial measurement  $M_0 = 200$  and the *fastest* measurement decrease  $\Delta_t = 100$  among the four settings in 4 - 7 rows of Table I achieves similar performance compared to the cases with a larger number of measurements  $M_0 > 200$  or  $\Delta_t > 100$ .

### C. Mask convergence

Finally, Fig. 10 compares the mask convergence of 1) UCB exploration in Section IV-C1, 2) decreasing the number of CS measurements combined with mask regularization in Section IV-C3, and 3) mask regularization only in Section IV-C, after 4000 time steps. We can observe that all of the three masks can capture the basic structure and the dominant directions of the AoD distribution in the beamspace. For UCB exploration in Fig 10.a, the mask fails to capture some *less likely* beam directions. The mask power allocation is more centered around several dominant directions, which makes it difficult to recover a beam direction that has not appeared too often as optimal. The mask convergence for linear measurement number decrease with  $M_0 = 300$ ,  $\Delta_M = 100$  is illustrated in Fig 10.b. As shown in Fig 10.b, the linear measurement decrease procedure achieves better exploration compared to UCB exploration in Fig 10.a. In addition, the mask is less *noisy* compared to the converged regularized mask in Fig 10.c. The number of CS measurements can be gradually

TABLE I  
COMPARISON OF THE AVERAGE BF LOSS OF MASK WITH PERFECT AOD PRIOR, UNIFORM MASK, THE MASK OBTAINED WITH UCB EXPLORATION, THE MASK ACHIEVED BY PROBABILITY ESTIMATION WITH MASK REGULARIZATION, AND THE MASK WITH CHANGING NUMBER OF MEASUREMENTS.

Avg. beamforming loss w.r.t opt. combining (dB) \ Number of measurements	$M = 25$	$M = 40$	$M = 50$	$M = 60$	$M = 80$	$M = 100$
2D-CCS with perfect AoD prior	3.90	2.73	2.33	2.07	1.62	1.40
2D-CCS with UCB exploration	8.16	4.49	3.83	3.47	2.63	3.06
2D-CCS with regularized mask	6.09	4.05	3.75	3.39	2.72	2.35
2D-CCS with uniform mask	11.18	6.73	5.86	5.31	4.29	3.63
Regularized mask with $M_0 = 200, \Delta_t = 100$	5.37	3.68	3.28	3.22	2.55	2.13
Regularized mask with $M_0 = 200, \Delta_t = 200$	5.18	3.67	3.26	3.03	2.52	2.01
Regularized mask with $M_0 = 300, \Delta_t = 100$	5.07	3.56	3.38	3.00	2.47	1.96
Regularized mask with $M_0 = 300, \Delta_t = 200$	5.21	3.62	3.41	3.13	2.50	2.05

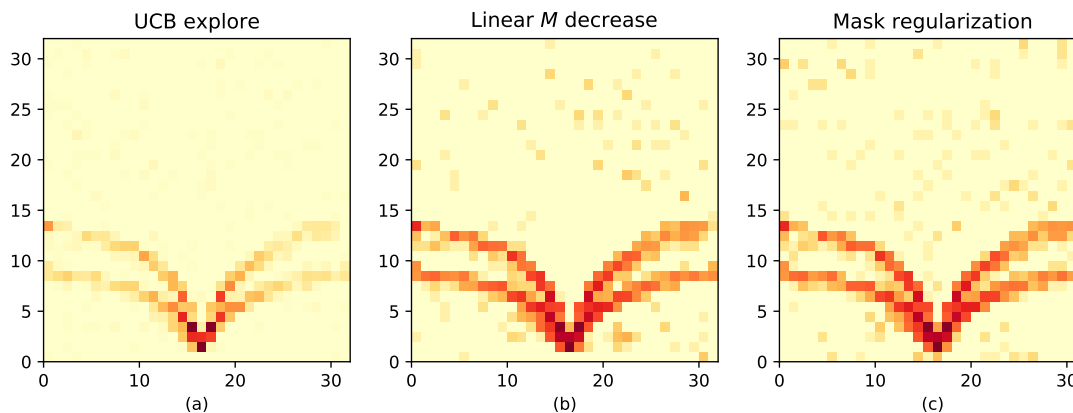


Fig. 10. Mask convergence for the online learning using UCB exploration, decreasing the number of CS measurements, mask regularization, with  $M$  (or  $M_{\min}$ ) = 80.

reduced with online learning. The prior learned with our online learning framework enables the BS to perform successful beam alignment with fewer channel measurements than standard CS.

## VI. CONCLUSION

In this paper, we proposed a site-specific online sensing matrix learning and beam alignment solution in vehicular communication. In a typical vehicular context, the AoD distribution can be structured due to the regularity of the vehicle trajectories and stationary objects such as buildings and roads. Leveraging such AoD distribution statistics in a vehicular scenario, we proposed a novel online compressive sensing framework using a 2D-CCS-based design. We formulated a convex optimization problem that uses the learned AoD prior to design a CS matrix that maximizes the approximate beam alignment probability. We observed superior performance of the AoD prior-aided sensing matrix design compared to either the standard 2D-CCS without AoD prior or exhaustive beam search. Using only 60 channel measurements, there was less than 2 dB gap between the average beam RSRP of the proposed approach and perfect CSI-aided optimal beamforming.

MmWave vehicular communication has specific channel statistics that need to be exploited for wireless system design. For future work, the design of the *subsampling set* in 2D-CCS that is well-suited to the AoD prior will be investigated. Another interesting topic is to leverage the potential angular statistics related to the receiver side for fast vehicle

beam alignment. As discussed in Section II-B, AoA prior in the vehicular context can be more complicated due to the highly mobile vehicles and varying traffic conditions. Side information such as vehicle locations or traffic conditions can be used as *context* to improve the online beam codebook optimization at the receiver side. For example, we can use contextual bandits-based online learning solution given appropriate side information. Furthermore, investigating other priors that are based on beam RSRP is an interesting research direction. In addition to the special street layout and specific channel distribution, more information is already embedded in *connected vehicles*. With side information such as location-based situational awareness, the number of measurements can further be reduced if the online learning can reveal the AoD distribution conditioned on the observed side information.

## REFERENCES

- [1] J. Choi, V. Va, N. Gonzalez-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath, "Millimeter-wave vehicular communication to support massive automotive sensing," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 160–167, 2016.
- [2] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal. Process.*, vol. 10, no. 3, pp. 436–453, 2016.
- [3] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.



- [4] T. S. Rappaport, R. W. Heath Jr, R. C. Daniels, and J. N. Murdock, *Millimeter wave wireless communications*. Pearson Education, 2014.
- [5] A. Festag, "Standards for vehicular communication from IEEE 802.11 p to 5G," *e & i Elektrotechnik und Informationstechnik*, vol. 132, no. 7, pp. 409–416, 2015.
- [6] S. A. A. Shah, E. Ahmed, M. Imran, and S. Zeadally, "5G for vehicular communications," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 111–117, 2018.
- [7] Z. Marzi, D. Ramasamy, and U. Madhow, "Compressive channel estimation and tracking for large arrays in mm-wave picocells," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 514–527, 2016.
- [8] A. Ali, N. González-Prelcic, and R. W. Heath, "Millimeter wave beam-selection using out-of-band spatial information," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1038–1052, 2017.
- [9] J. Santa and A. F. Gomez-Skarmeta, "Sharing context-aware road and safety information," *IEEE Pervasive Comput.*, vol. 8, no. 3, pp. 58–65, 2009.
- [10] P. Papadimitratos, A. L. Fortelle, K. Evenssen, R. Brignolo, and S. Cosenza, "Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation," *IEEE Commun. Mag.*, vol. 47, no. 11, pp. 84–95, 2009.
- [11] V. Va, J. Choi, T. Shimizu, G. Bansal, and R. W. Heath, "Inverse multipath fingerprinting for millimeter wave V2I beam alignment," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4042–4058, 2017.
- [12] Y. Wang, A. Klautau, M. Ribero, A. C. Soong, and R. W. Heath, "MmWave vehicular beam training with situational awareness using machine learning," *IEEE Access*, 2019.
- [13] N. González-Prelcic, A. Ali, V. Va, and R. W. Heath, "Millimeter-wave communication with out-of-band information," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 140–146, 2017.
- [14] A. Ali, N. González-Prelcic, and A. Ghosh, "Passive radar at the roadside unit to configure millimeter wave vehicle-to-infrastructure links," *arXiv preprint arXiv:1910.10817*, 2019.
- [15] —, "Millimeter wave V2I beam-training using base-station mounted radar," in *Proc. Radar Conf.* IEEE, 2019, pp. 1–5.
- [16] N. González-Prelcic, R. Méndez-Rial, and R. W. Heath, "Radar aided beam alignment in mmWave V2I communications supporting antenna diversity," in *Proc. Inf. Theory Appl. Workshop (ITA)*. IEEE, 2016, pp. 1–7.
- [17] A. Klautau, N. González-Prelcic, and R. W. Heath, "LIDAR data for deep learning-based mmWave beam-selection," *IEEE Wireless Commun. Lett.*, 2019.
- [18] N. Garcia, H. Wymeersch, E. G. Ström, and D. Slock, "Location-aided mm-wave channel estimation for vehicular communication," in *Proc. Int. Workshop Signal Process Advances Wireless Commun. (SPAWC)*. IEEE, 2016, pp. 1–5.
- [19] Y. Wang, K. Venugopal, A. F. Molisch, and R. W. Heath, "MmWave vehicle-to-infrastructure communication: Analysis of urban microcellular networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7086–7100, 2018.
- [20] A. Klautau, P. Batista, N. Gonzalez-Prelcic, Y. Wang, and R. W. Heath, "5G MIMO data for machine learning: Application to beam-selection using deep learning," in *Proc. Inf. Theory Appl. Workshop (ITA)*. IEEE, 2018, pp. 1–9.
- [21] V. Va, T. Shimizu, G. Bansal, and R. W. Heath, "Online learning for position-aided millimeter wave beam training," *IEEE Access*, vol. 7, pp. 30 507–30 526, 2019.
- [22] S. Shalev-Shwartz *et al.*, "Online learning and online convex optimization," *Foundations and Trends® in Machine Learning*, vol. 4, no. 2, pp. 107–194, 2012.
- [23] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *J. Machine Learning Research*, vol. 11, no. Jan, pp. 19–60, 2010.
- [24] J. Kivinen, A. J. Smola, and R. C. Williamson, "Online learning with kernels," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2165–2176, 2004.
- [25] G. H. Sim, S. Klos, A. Asadi, A. Klein, and M. Hollick, "An online context-aware machine learning algorithm for 5G mmWave vehicular communications," *IEEE/ACM Trans. Network.*, vol. 26, no. 6, pp. 2487–2500, 2018.
- [26] M. Hashemi, A. Sabharwal, C. E. Koksal, and N. B. Shroff, "Efficient beam alignment in millimeter wave systems using contextual bandits," in *Proc. Conf. Comput. Commun.* IEEE, 2018, pp. 2393–2401.
- [27] J. Gittins, K. Glazebrook, and R. Weber, *Multi-armed bandit allocation indices*. John Wiley & Sons, 2011.
- [28] A. Garivier and E. Moulines, "On upper-confidence bound policies for switching bandit problems," in *Int. Conf. Alg. Learning Theory*. Springer, 2011, pp. 174–188.
- [29] S. Ji, Y. Xue, L. Carin *et al.*, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, p. 2346, 2008.
- [30] A. C. Kak, M. Slaney, and G. Wang, "Principles of computerized tomographic imaging," *Medical Physics*, vol. 29, no. 1, pp. 107–107, 2002.
- [31] M. Hassan-Ali and K. Pahlavan, "A new statistical model for site-specific indoor radio propagation prediction based on geometric optics and geometric probability," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 112–124, 2002.
- [32] S. Y. Seidel and T. S. Rappaport, "Site-specific propagation prediction for wireless in-building personal communication system design," *IEEE Trans. Veh. Technol.*, vol. 43, no. 4, pp. 879–891, 1994.
- [33] Y.-G. Lim, Y. J. Cho, Y. Kim, and C.-B. Chae, "Map-based millimeter-wave channel models: An overview, guidelines, and data," *arXiv preprint arXiv:1711.09052*, 2017.
- [34] N. J. Myers, A. Mezghani, and R. W. Heath Jr, "FALP: Fast beam alignment in mmWave systems with low-resolution phase shifters," *arXiv preprint arXiv:1902.05714*, 2019.
- [35] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, 2013.
- [36] A. C. Kak, M. Slaney, and G. Wang, "Principles of computerized tomographic imaging," *Medical Physics*, vol. 29, no. 1, pp. 107–107, 2002.
- [37] H. O. Lancaster and E. Seneta, "Chi-square distribution," *Encyclopedia of Biostatistics*, vol. 2, 2005.
- [38] S. Diamond and S. Boyd, "CVXPY: A Python-embedded modeling language for convex optimization," *The J. Mach. Learning Research*, vol. 17, no. 1, pp. 2909–2913, 2016.
- [39] R. W. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.
- [40] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multi-armed bandit problem," *Machine learning*, vol. 47, no. 2-3, pp. 235–256, 2002.
- [41] T. Lattimore and C. Szepesvári, *Bandit Algorithms*. Reading, MA: Addison-Wesley, 1972.
- [42] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [43] P. C. Mahalanobis, "On the generalized distance in statistics." National Institute of Science of India, 1936.
- [44] R. Beran *et al.*, "Minimum hellinger distance estimates for parametric models," *The annals of Statistics*, vol. 5, no. 3, pp. 445–463, 1977.